Quality assessment of digital soil maps: producers and users perspectives.

Peter A. Finke
Biometris, Wageningen University

Abstract
The assessment of quality of soil maps can be seen from the producers and the users perspective. Producers perspectives have led to several measures of accuracy and precision that describe the intrinsic quality of the produces soil map and information system. These are described in some detail. In conclusion, it seems that adequate measures lack to detect and quantify logical inconsistencies as resulting from joining and harmonisation of existing maps. Additionally, indicators need to be developed that assess the semantic quality of maps while accounting for the taxonomic distance between the map units. Users perspectives lead to a different view on map quality. Some minimal data sets of error are proposed that will enable users to incorporate soil map uncertainty into their applications.

1. Introduction

More than half a century of soil inventories has resulted in a great amount of soil data sets, collected and presented in various ways via maps and soil information systems. During this period, mapping scales varied, mapping methods changed and mapped areas increased. The resulting conglomerate of soil information systems is being used for a wide range of applications. This opens the question of quality assessment, since ignoring uncertainties associated with the soil data sets in interpretations, may lead to wrong decisions and will also reduce the confidence in soil scientists (Fisher, 1999).

This recognition has lead to studies on the issue of uncertainty in the context of geographical information systems. Zhang and Goodchild (2002) used the term uncertainty as an umbrella for the distinct terms error, randomness and vagueness. The spatio-temporal soil system that we try to understand and describe during the soil mapping process, is associated with randomness, if only because we do not fully understand it and cannot completely measure it. The descriptive model for this system (i.e., the combination of the data acquisition, the derived conceptual model and the applied map inference methods) is associated with error and vagueness, depending on the type of chosen model and its methods.

Quality in relation to cartography has been defined as (Moellering, 1987) *the suitability of the data for the intended use*. Others (e.g., Forbes et al., 1982) use terms like *adequacy* and *fitness for use* to describe the quality of soil resource inventories. These approaches have in common that quality is made dependent on the intended use of the inventoried data, which is seldom single-purpose. Any usage of soil resource data is associated with a characteristic sensitivity of the application to variation, due to errors in these data. The same error in basic soil data will cause different uncertainties in different applications (e.g., Finke et al., 1996), and thus the usage aspect of quality is in practice a variable. One objective of this paper is therefore to indicate some methods to assess the impact of the quality of soil mapping to applications.

Other aspects of quality are constant for the inventory considered because they directly apply to the data. Since 2002, an ISO standard for describing the quality of geographic data exists (ISO 19133), based on a conceptualisation of clear and identifiable objects. Fisher (2003; Table 1) introduced similarities between uncertainty terms and data quality components, while stressing the limitations of the ISO-standard for indeterminate objects. Apparently, the list of data quality components does not reflect all perceived modes of uncertainty and vice versa. Furthermore, indicators of quality are often associated with the methods applied to map soils. Another objective of this paper is therefore to give some quality measures associated to different methods of soil mapping and to different aspects of quality.

In comparison to traditional soil mapping methods, modern predictive methods are better defined, documented and thus less depent on the individual surveyor’s style. Also, modern methods come often with an indication of quality. This leads to two questions that deserve elaboration:

1. Do we expect modern predictive soil maps to be more accurate than conventional soil maps?
2. Do we need new ways of assessing quality?

2. Assessment of intrinsic quality of soil resource data

2.1 Notions and definitions

Throughout this paper the notions of *accuracy* and *precision* will be used. The definitions, put in the context of soil mapping, are (after Burrough, 1990):

Accuracy: The degree of conformity of the soil map with reality;
Precision: The measure of uncertainty associated with the set of procedures used to map the reality.

Accuracy measurements are taken to (in-) validate the soil map and should be independently collected. Precision measures often come as a by-product from the (predictive) soil mapping method and give some kind of prediction error. Below I will summarise, with reference, some accuracy and precision measures for the components of soil data quality of Table 1.

2.2 Positional Quality

Positional quality refers to several aspects of the geo-referencing and topology of the soil map. A distinction is made between the positional quality that is reflected in the intricacy of the geographic patterns and the one that is associated with the positioning and widths of individual soil boundaries.

2.2.1 Effective map scale

The effective map scale is a precision measure that demonstrates if the intricacy of the patterns depicted on the soil map corresponds to the presentation scale. The effective scale number *ESN* of a map extent is calculated as (after Forbes et al., 1982):

$$ESN = NSN \times \sqrt{\frac{\sum_{j=1}^{m} A_j}{m \times MLD \times 4}} \qquad \text{-1-}$$

where A_j is the area of the j th polygon and m the total number of (complete, soil) polygons within the map extent, *MLD* is the factor by which the scale of the map could be reduced before the average polygon area would equal the smallest legible area on the map (*MLD* is usually set to 0.4 cm^2), and *NSN* is the nominal scale number. The number "4" certifies that the average polygon area is four times the MLD when the presentation scale is equal to the effective scale.

2.2.2 Location and width of soil boundaries

The location of map boundaries is associated with uncertainty because of positional error, digitising error and the artificiality of many boundaries. The uncertainty with respect to the location of soil boundaries can be expressed in various ways. In case ground truth is available (a control soil mapping with the same map legend but at a more detailed scale), the accuracy measure Area of Disagreement (*AD*) can be counted, which is the intersection between the coverage of the evaluated map and the control map. Such control soil mapping may be a physiographic landform map derived from DEM (Hengl, 2003, p.171), if it is certain that all soil polygon boundaries should be physiographic boundaries as well. The positional accuracy of the delineation (PA_d) can then be calculated as:

$$PA_d = \frac{AD}{l + l'} \qquad \text{-2-}$$

where l is the boundary length of the delineation on the soil map, and l' is the boundary length of the delineation on the control map.

Vice versa, if an estimate of the positional accuracy is available, e.g., because it has been measured in transects, so called Epsilon bands (Chrisman, 1982) can be constructed to indicate areas around boundaries that are uncertain. The summed-up areas of the epsilon bands as percentage of the map area then indicate the *AD*.

Ideally, mapped soil boundaries are situated at the location where the highest rate of change occurs (Burrough, 1990) in those soil properties that define the soil map unit. The width of the boundary could then be defined as the distance along the gradient of change where a predefined minimal rate of change is exceeded. In case a fuzzy soil map is available, a precision measure for the boundary width of the derived defuzzified soil map units can be calculated using the membership gradients of the fuzzy classes on each side of the border of the defuzzified map. Accuracy can be calculated along densely sampled transects. In case of traditional soil mapping, Clarke et al. (1991) propose recording the estimated boundary width and the rank of the importance of the characteristics used to determine the boundary in the field during the mapping.

2.3 Attribute Quality

The quality of data in soil databases is determined by errors related to the measurement, such as the method of analysis and the laboratory. But it may also depend on the currency of the measurements, the sampled and analysed volumes and errors introduced during digitisation. Uncertainty is usually estimated by some RMSE-measure, sometimes specified for specific laboratories or methods of analysis. Values may be high (Van Reeuwijk and Houba, 1998) but are often undocumented in soil databases. Estimation of accuracy is not easy since true validation samples (with known values) are almost inevitably artificial and values in databases relate to natural samples. Storing data on the date of measurement, analysis protocol and laboratory along with the analysed data provides at least a framework for the assessment of precision of the data.

2.4 Completeness

The quality of a soil database is, from the perspective of a user, often determined by its completeness, the degree to which the necessary data are present. Many soil databases suffer from unsatisfactory completeness, both geographically (data density relative to the map scale) and thematically (attribute completeness).

If there is a clear definition of "necessary data" (which is application dependent), the attribute completeness S can simply be counted as an area-weighted data saturation fraction:

$$S = \sum_{i=1}^{n} W_i \frac{r_{data,i}}{r_{data,i} + r_{mis,i}} \qquad \text{-3-}$$

where W_i is the areal fraction assigned to the ith sample location (e.g., by spatial declustering; Dubois and Saisana, 2002), r_{data} is the number of database fields filled with data and r_{mis} is the number of data base fields containing missing data.

Part of attribute incompleteness can be resolved by estimation through continuous, class or taxonomic pedotransfer functions (Wösten et al., 1995; Van Ranst et al., 1995). Nevertheless, these approaches often lack an indication of the accuracy of the estimated values, and are often targeted towards less easy to measure parameters using basic soil data in the estimation process. To improve the completeness of the basic soil data itself, data imputation techniques are available that give an estimation error as well. Reference is made to Cohen (1996) and Rubin (1987) for an overview of data imputation techniques in monovariate and multivariate cases. Given the current availability of interpolation methods that can utilise "soft data", there is no reason to discard incomplete profile data from predictive soil mapping. For the cases with censored data in the database ("deeper than", "more than"), methods are available to make (e.g., maximum likelihood) estimates (Cohen, 1991; Knotters et al., 1995).

2.5 Semantic Quality

The semantic quality of a soil map can be described as the degree of reliability of the map legend when field-checked, but also as the degree the map legend maximally mirrors the natural variation. The characteristics that identify the legend entity should lie in the appropriate range or have the right (thematic) value when checked. Accuracy and precision parameters relating to identifying characteristics are described in the next sections. In case of thematic maps, the degree of separation between the identifying characteristics between mapping units also defines its semantic quality (e.g., Webster and Oliver, 1990), but this will not be treated here.

Other factors, related to descriptive characteristics, are here considered not to contribute to semantic quality. More specifically, the degrees to which legend entities are mutually different and internally homogenous with respect to descriptive characteristics do not describe the intrinsic quality of the legend. Nevertheless, these factors may dominate the user quality, since SMU resulting from a general-purpose mapping may be sub-optimal to serve as geographic building blocks units for specific applications if these SMU do poorly explain the deviance of the descriptive characteristics.

2.5.1 Accuracy of thematic maps

The quality of thematic soil maps can be assessed by a comparison of field observations with predictions by the soil map. Such validation results for one soil class at the observation level in either a true or a false, and thus the binomial distribution can be used for testing. Ideally, a sample is taken so that all occurring soil map classes are visited. The result is stored in a so-called confusion matrix (Lillesand and Kiefer, 1994), which is the basis for multinomial testing. A simple example whereby all samples are weighted equally is given in Table 3. Weighting can be applied to account for errors that are considered less important, e.g. because the two classes involved are considered to be taxonomically adjacent. There is no agreement on how such taxonomic weighting should be done.

The confusion matrix can be used to calculate a number of statistics (symbols are introduced in Table 3, values range between 0=poor and 1=excellent):

Taxonomic purity or over-all accuracy: $\theta_1 = \sum_{i=1}^{c} p_{ii}$; -4-

Users' accuracy for class i: p_{ii} / p_{i+} -5-

Producers' reliability for class *j*: p_{jj} / p_{+j}

The remote sensing community has developed some alternative statistics, basically with the purpose to motivate choices between classification routines. These may be applicable to soil mapping since classification of remote sensing and digital elevation data plays an important role in modern soil mapping methods.

Some of the classification accuracy presented in the confusion matrix may be due to chance, because some classes may occupy much larger areas than others and thus dominate the validation sample. For these circumstances, the *kappa* statistic (Cohen, 1968) has been developed. *Kappa* (< 0: the map performs worse than randomly distributed classes; 1=excellent) is calculated by:

$$\kappa = \frac{\theta_1 - \theta_2}{1 - \theta_2} \qquad \text{-6-}$$

where

$$\theta_2 = \sum_{i=1}^{c} p_{i+} p_{+i} \qquad \text{-7-}$$

When full-cover mapping is not yet done but the validation sample is already taken, instead of *kappa*, the *Tau* statistic (Ma and Redmond, 1995; ranges comparable to those of *kappa*) can be calculated using prior probabilities p_i of class memberships (e.g. estimated from small-scale soil maps or terrain maps):

$$\tau = \frac{\theta_1 - \theta_2^{'}}{1 - \theta_2^{'}} \qquad \text{-8-}$$

where

$$\theta_2^{'} = \sum_{i=1}^{c} p_i p_{+i} \qquad \text{-9-}$$

2.5.2 Precision of fuzzy maps

The maps of the memberships of fuzzy classes can be used to calculate the confusion index (*CI*) at all map pixels or of the map extent, which is a precision measure calculated by the ratio of the first and second membership *m*:

$$CI = m_{2nd\max} / m_{1st\max} \qquad \text{-10-}$$

CI varies between 0 (no confusion) and 1 (maximal confusion), and can be used to indicate transition zones. A weakness of this index is that is accounts only for the membership confusion and not for the taxonomic distance between the two fuzzy classes. Hengl et al. (2004) recently proposed a method to combine taxonomic distance by colour separation and confusion by whiteness saturation into one map colour coding system.

2.5.3 Quality of single value maps

The quality of single value maps can be expressed by the accuracy measure MSE (Mean Square Error) and by comparable precision measures such as PEV (Prediction Error Variance) associated with the prediction method (Webster and Oliver, 1990). Additionally, the percentage of variance explained can be calculated and the conditional bias (second regression) can be determined. Since soil mapping does usually not lead to single value maps, associated quality measures are not extensively treated.

2.6 Currency

The currency of a soil map is a function of its age, because the soil system that it describes or the concepts and methods that are used for the description of the soil system may have changed. An outdated map needs to be updated but may keep its value as an historical document, e.g. to assess historical carbon stocks. The degree of ageing can in some instances be monitored using measures of positional or semantic quality. For example, Finke (2000) gives two parameters to assess map quality of ground water table class maps in The Netherlands. These parameters are both based on point values of a function G that describes the degree to which two ground water fluctuation parameters deviate from the definition in the map legend in a particular year. The first parameter is the estimated average value of G for a map sheet:

$$MG = \sum_{i=1}^{n} g_i \cdot G_i \qquad \text{-11-}$$

where g_i is the weight assigned to a point value of G, depending on the sampling design, and all weights g_i sum up to 1.

The second parameter is the estimated fraction of the area with strong deviations from the map legend:

$$FEXG = \sum_{i=1}^{n} g_i \cdot I_i \qquad \text{-12-}$$

with I_i=1 if G_i≥1 and I_i=0 if G_i<1.

MG and *FEXG* can be monitored through *G*. In Figure 1 the evolution of MG and FEXG for one map sheet is shown. The graph also shows empirically derived threshold quality values to support decisions on map updating.

2.7 Logical consistency

Logical consistency of a soil map means, that no interpretative mistakes have been made in the mapping process that are reflected in the final maps. Interpretative mistakes may (e.g.) occur during field mapping, generalisation, combination and harmonisation of adjacent maps. It becomes more important to assess this type of error when soil information systems obtain a more composite nature, as they are developed out of regional existing information systems (Van Engelen and Wen, 1995; Deckers et al., 1998; Finke et al, 2003; Lambert at al., 2003).

An example of an inconsistency due to generalisation error is the situation where one Soil Map Unit (SMU) at the detailed scale is assigned to 2 SMU that share a boundary at the generalised scale (Figure 2).

Logical inconsistencies may occur as well when results of several mapping projects are combined. An example is given in Figure 3, where national boundaries are visible in trans-national soil maps (Lambert at al., 2003; European Soil Bureau Network, 2004). These inconsistencies may occur for conventional soil maps, where different surveyors may have taken different classification or delineation decisions in comparable field situations. It may occur as well in pedometric mapping, when clustering methods (e.g., resulting in membership maps derived from fuzzy-*k*-means classification) have been applied, because when training data sets differ, so may the resulting classifications.

This type of logical inconsistency may be recognised visually in thematic maps, but is not easily automatically detected or quantified via precision or accuracy measures. A proposed approach is to use a full-coverage landform classification based on a DEM (the classification possibly being supervised with Aerial Photo Interpretation; Hengl and Rossiter, 2003) to detect inconsistencies in the combined soil maps:

(i) The Physiographic Units (PU) are identified and mapped over the full map extent;

(ii) The PU are split up by the SMU boundaries in each one soil map that contributes to the combined soil map, through the operation PU2=PU∩SMU. The associated boundary uncertainty can be estimated through identifying the Area of Disagreement (AD), a positional accuracy measure. Alternatively, Epsilon Bands, a positional precision measure, are constructed and their area EBA is counted (section 2.2).

(iii) The AD or EBA polygons are removed from the PU2 trough the operation PU3=PU2-(PU2∪[AD,EBA]) and the resulting polygons PU3 are joined at the map boundaries.

(iv) After the combination of the soil maps, those polygons are selected from PU3 (PU4⊆PU3), that are again split up due to the SMU boundaries of the combined soil maps. These polygons PU4 are suspect. The precision measure Area of Logical Inconsistency (ALI) is then estimated by:

$$ALI = \sum_{i=1}^{n} Area(PU4_i) \qquad \text{-13-}$$

2.8 Lineage

Lineage is a possible source of uncertainty, because it may involve that errors are introduced when integrating data from different sources, possibly of various ages. A well-known example is the lesser quality of positional data in older topographic maps and information systems, which influences the positional quality of (parts of) the soil map (e.g., Radoševic, 1979). Another example is the usage of recent, detailed soil maps to update parts of smaller-scale soil maps (e.g., Finke et al., 2004). The adaptation of updates from one data set to another leads to the problem of integration of heterogeneous data. If the data model of the different data sets is homogeneous, integration of the data is of a geometric nature. Else, a semantic integration must be done first to avoid faulty comparisons and subsequent logical inconsistencies. Walter and Frisch (1996) evaluate some statistical approaches towards this type of data integration and associated precision measures, but few methods to quantify the uncertainty effect of lineage seem to be existing.

3. User Quality

3.1 Attitudes towards uncertainty

Uncertainty often causes dilemmas for the people who are exposed to it, especially when the uncertain data are to be used to support (the development of) policy. Policy makers and stakeholders then become users of uncertain data (and its interpretations). As such, they may experience both advantages (room for improvement, a bandwidth for making decisions and room for argumentation) and disadvantages (loss of public image and the risk to make wrong decisions) from uncertain data. All three types of users benefit by reduction of the disadvantages of uncertainty in the data (Table 2). It is therefore safe to assume that there will be broad support for activities reducing the disadvantages associated with uncertainty for all three types of users. Quantifying uncertainty is one of these activities, when it is done in such way, that it allows for the identification of the sources of uncertainty to be able to minimise uncertainty. This requests interaction between the data collectors and the data users (Figure 4). Also, uncertainty should be quantified in such manner, that it is useable for methods of decision making (e.g., Raiffa and Schlaifer, 2000). Finally, there is the issue of how to communicate uncertainty and associated risk to stakeholders (e.g., Gutteling and Wiegman, 1996) but this is considered beyond the scope of this paper.

In the following sections, some examples will be given on the utilisation of uncertain soil- and landscape data in policy support studies. Focus will be on the description of uncertainty in an error model relevant to the application. Some of the example studies take the form of an uncertainty analysis. The necessary error model can be considered as the “minimum data set of uncertainty” for these example studies and may differ considerably from the intrinsic data quality measures described earlier as they serve a different purpose.

3.2 An error model for evaluation studies using crisp thematic maps

This example is taken from a study by Finke et al. (1999), in which the effect of errors in categorical data (i.e., the generalised soil and vegetation class maps of the EU) on the uncertainty of outcomes of a soil acidification model was analysed. A deposition scenario from the Netherlands environmental outlook (RIVM, 1997) was simulated. To assess the quality of the EU maps, highly detailed maps of soil and vegetation available for the Netherlands (NL) were used as ground truth.

To quantify the degree of error within each EU-category, an indicator variable $I_{t,s}$ for NL soil-vegetation class t within EU soil-vegetation class s was introduced:

$$I_{t,s}(x) = \begin{cases} 1 & \text{if } \mathrm{EU}(x) = s \text{ and } \mathrm{NL}(x) = t \\ 0 & \text{otherwise} \end{cases} \qquad \text{-14-}$$

where EU(x) is the soil and vegetation class at location x in the EU map and NL(x) is the ground truth according to the detailed map. The error model was defined as:

a. The confusion matrix (Lillesand and Kiefer, 1994) with the expectations of I for all combinations of t and s;
b. The binominal variances $s^2(I_{t,s}(x))$ associated with the cells in the confusion matrix;
c. Fitted indicator variograms $\gamma_{I_{t,s}}(h)$, scaled so that the sill equals the binominal variance, for $I_{t,s}(x)$ for all non-zero expectations of I in the confusion matrix. The incorporation of the spatial correlation of the error stemmed from the observation that misclassifications tend to appear in clusters.

The analysis was part of an uncertainty analysis, which also included the effect of uncertainty in continuous data. Because the output of the involved model was known to respond non-linearly to its inputs, the uncertainty analysis was set up as a Monte Carlo analysis. This approach requires the generation of realisations of model inputs and the error model should allow for this to be done effectively. Thus, this study, reported in Kros et al. (1999), consisted of the following steps:

1. 25 Realisations of the EU-map were obtained by sequential multiple indicator simulation (Goovaerts, 1997), using the indicator variograms and a stratification to EU-categories.
2. For each one map realisation, 5 realisations of soil parameters and 5 realisations of vegetation parameters were simulated using nonconditional sequential multivariable Gaussian simulation (Pebesma and Wesseling, 1997). For the error model of these continuous data, reference is made to section 3.4.
3. The acidification model was run on all 625 input data sets.

4. The output uncertainty and the relative contributions caused by uncertainty in categorical data, soil parameters, vegetation parameters and an interaction term were quantified in a (nested) ANOVA.

3.3 An error model for evaluation studies using fuzzy thematic maps

This example is taken from a study by Gorsevski et al. (2003), in which a continuous landform classification by fuzzy *k*-means is combined with a Bayesian probabilistic modelling approach to obtain probabilistic landslide hazard maps.

The Bayesian approach combines subjective probability with conditional probability. The subjective probability expresses the degree of belief in an event (i.e. the probability of occurrence of a landslide) and is usually called the *prior*. The conditional probability expresses the likelihood of the hypothesis to be true given the evidence (i.e. the probability of occurrence of a landslide given the fuzzy memberships). To calculate the probability of occurrence of a landslide at a location, the following equation applies:

$$P(o \mid f) = \frac{P(f \mid o) \cdot P(o)}{P(f)} \text{ (Bayes' rule)} \qquad \text{-15-}$$

where *o* indicates the occurrence of a landslide and *f* the set of memberships associated with the fuzzy clusters. *P(o)* is the prior probability for occurrence of a landslide, estimated by the counted occurrences of landslides. *P(f|o)* is the conditional probability of occurrence of a landslide given the fuzzy memberships, calculated from the relative frequency of association between occurrences of landslide locations and categorised membership values of the fuzzy *k*-means classes at the *n* locations (grid cells) for which the occurrence (or absence) of landslides is known:

$$P(f \mid o) = \prod_{i=1}^{n} co_i \qquad \text{-16-}$$

where $\prod_{i=1}^{n} co_i$ is the product of conditional probabilities for occurrence for attributes *i*=1..*n* of the predictor data sets. *P(f)* is calculated by

$$P(f) = P(f \mid o) \cdot P(o) + P(f \mid a) \cdot P(a) \qquad \text{-17-}$$

where *a* indicates the absence of a landslide and *P(f|a)* and *P(a)* are calculated analogously to *P(f|o)* and *P(o)* respectively.

Combination of the above three equations leads to the equation for the Bayesian calculation:

$$P(o \mid f) = \frac{P(o)\prod_{i=1}^{n} co_i}{P(o)\prod_{i=1}^{n} co_i + P(a)\prod_{i=1}^{n} ca_i} \qquad \text{-18-}$$

The analyses comprised the following steps:

1. Translation of the DEM into landslide-relevant environmental attributes and performing a fuzzy *k*-means classification on training areas. Using the performance indicator *FPI* (Minasny and McBratney, 2000), the optimal number of classes was found; with this number of classes, the optimal fuzzy exponent ϕ was derived after Odeh et al. (1992) and the memberships at all grid cells of the DEM were calculated.
2. Recording the occurrences of landslide as absence or presence at all grid cells of the DEM.
3. Construction of an error model containing:
   a. Tabulated conditional probabilities of occurrence and absence (expressed as relative frequencies) for each one fuzzy cluster (subdivided in 10 membership subclasses 0-0.1, 0.1-0.2, … 0.9-1.0).
   b. Prior probabilities of absence and occurrence of landslides
   c. Maps of the memberships of each one fuzzy class.
4. Calculation at each one map pixel of the conditional probabilities based on components a and b of the error model.
5. Calculation of the Bayesian probability of occurrence of landslides at each one map pixel, using equation 18.

3.4 An error model for evaluation studies using (multiple) single value maps

This example is taken from the same study by Kros et al. (1999) that was briefly described in section 3.3. Below, I focus on step 2 of the over-all procedure described in section 3.3. A commonly applied method to generate realisations of (multiple) single value maps (Gómez-Hernández and

Journel, 1992) is that of joint sequential simulation of Gaussian fields. The necessary error model includes:

a. For each parameter, possibly per stratum, the average value and the variogram.
b. Between parameters, possibly per stratum, the cross-variograms.

To reduce the calculation effort, before starting the simulations, the sensitive parameters were identified. Further simplifications comprised the reduction of the number of parameters that were supposedly correlated, and the assumption that cross-variograms would lack spatial structure so that the covariance would be a constant.

4. Conclusive remarks

Question 1: Do we expect modern predictive soil maps to be more accurate than conventional soil maps?

- Yes, because the positional quality, currency and lineage of the data used to construct digital soil maps is better than that used to construct the conventional soil maps in the past. Additionally, the fact that precision measures come as by-products of many modern mapping methods allows for the optimisation of these methods. The circumstances are there to make better maps with less field effort, all though this advantage can be lost when too is much economised on the collection of ground truth data. The user of modern predictive soil maps may conclude from the simple presence of precision measures that quality of modern soil maps is less than that of conventional soil maps (because no quality indications were given with these maps). The importance of this aspect should not be neglected and requires communication to the users on quality aspects and on how to manage uncertainty in soil map applications.

Question 2: Do we need new ways of assessing quality?

- Work is needed on the assessment and improvement of uncertainty introduced by combining data from different surveys. There is a need to develop a framework for the harmonisation of soil maps.
- The assessment of semantic quality of soil maps can be improved by including the taxonomic distance as weight in confusion matrices, such that misclassifications over taxonomically adjacent classes is less weighted.
- The description of intrinsic quality is good for documenting and improving quality but may be insufficient to assess the user quality. Some standardisation in the description of quality in terms of uncertainty or error models is necessary, because it will help bridge the gap between data providers and data users.

This paper has focused on the assessment of several aspects of quality, and on error models to utilise quality for users. It may leave the impression that modern mapping methods may improve quality while conventional soil maps will pass into oblivion. This is certainly not the case. Modern methods may and should be used to improve the quality of existent soil information systems in three ways:

(1) updating (improving the currency);
(2) upgrading (improving the completeness);
(3) Corroboration (improving the positional and semantic quality and the logical consistency).

Some experience with (1) and (2) exists (e.g., Finke et al., 2004). The toolbox for (3) is already well filled but applications seem yet to be absent.

References

Burrough, P.A. 1990. Principles of Geographical Information Systems for Land Resource Assessment. Clarendon Press. Oxford.

Burrough, P.A., P.F.M. van Gaans and R. Hootsmans, 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. Geoderma 77: 115-135.

Chrisman, N. R. (1982). A theory of cartographic error and its measurement in digital data bases. Proceedings, Auto Carto 5, 159-168.

Clarke, S.E., D. White, and A.L. Schaedel. 1991. Oregon, USA ecological regions and subregions for water quality management. Environmental Management Vol. 15, No. 6, pp. 847-856.

Cohen, J., 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70:426–443.

Cohen, A.C., 1991. Truncated and censored samples: theory and applications. Dekker Inc., New York.

Cohen, M.P. ,1996. A new approach to imputation. Proceedings of the Survey Research Methods Section of the American Statistical Association, 293-298.

Deckers, J.A., F.O. Nachtergaele and O.C. Spaargaren (eds.), 1998. World Reference Base for soil resources. Introduction. 84 World Soil Resources Report, FAO, ISRIC and ISSS. Rome, Italy.

Dubois G. and M. Saisana. 2002. Optimizing Spatial Declustering Weights – Comparison of Methods. Proceedings of the Annual Conference of the International Association for Mathematical Geology, Berlin-Germany, September 15-20, 2002, p. 479-484.

European Soil Bureau Network. 2004. Soil Map Internet Server. http://eusoils.jrc.it/msapps/Soil/SoilDB/SoilDB.phtml (accessed 19 june 2004).

Finke, P.A., J.H.M. Wösten and M.J.W. Jansen. 1996. Effects of uncertainty in major input variables on simulated functional soil behaviour. Hydrological Processes 10, 661-669.

Finke, P.A., D. Wladis, J. Kros, E.J. Pebesma and G.J. Reinds, 1999. Quantification and simulation of errors in categorical data for uncertainty analysis of soil acidification modelling. Geoderma 93: 177-194.

Finke, P.A. 2000. Updating groundwater table class maps 1:50,000 by statistical methods: an analysis of quality versus cost. Geoderma 97: 329-350.

Finke, P.A., R. Hartwich, R. Dudal, J. Ibanez, M. Jamagne, D. King, L. Montanarella and N. Yassoglou. 2003. Geo-referenced soil database for Europe. Manual of procedures, version 1.1. EUR 18092 EN European Soil Bureau, JRC, Italy.

Finke, P.A., M.F.P. Bierkens, D.J. Brus, T. Hoogland, M. Knotters, F. de Vries. 2004. Mapping ground water dynamics using multiple sources of exhaustive high resolution data. Geoderma, in press.

Fisher, P. F. 1999. Models of uncertainty in spatial data. In Longley et. al.: Geographical Information Systems, Volume 1: Principles and Technical Issues. John Wiley & Sons. ISBN: 0471331325.

Fisher, P. F. 2003. Data Quality and Uncertainty: Ships passing in the night! In Shi, W., Goodchild, M. F., Fisher, P. F.: Proceedings of ISSDQ 2003, Hong Kong. 13.

Forbes, T.R., D. Rossiter and A. van Wambeke, 1982. Guidelines for evaluating the adequacy of soil resource inventories. 1987 printing edition of SMSS Technical Monograph 4. Cornell University, Dept Agronomy, Ithaca, NY.

Gómez-Hernández, J.J. and A. G. Journel, 1992. Joint sequential simulation of multigaussian fields. In: A. Soares, editor, Geostatistics Troia '92, I., pages 85:94.

Gorssevski, P.V., P.E. Gessler and P. Jankowski, 2003. Integrating a fuzzy k-means classification and a Bayesian approach for spatial prediction of landslide hazard. J Geograph Syst 5: 223-251.

Goovaerts, P., 1997. Geostatistics for national resources evaluation. Oxford Univ. Press, 483 p.

Gutteling, J.M. and O. Wiegman, 1996. Exploring risk communication. Kluwer Academic Publishers, Dordrecht. 236 pp.

Hengl, T. 2003. Pedometric mapping. PhD-thesis, ITC, Enschede.

Hengl, T. and D.G. Rossiter, 2003. Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. Soil Sci Am J. 16(6): pp.1810-1822.

Hengl, T., D.J.J. Walvoort, A. Brown and D.G. Rossiter, 2004. A double continuous approach to visualisation and analysis of categorical maps. Int Jou Geoinf Sci 18 (in press).

Knotters, A., D.J. Brus and J.H. Oude Voshaar, 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. geoderma 67: 227-246.

Lambert, J.J., J. Daroussin, M. Eimberck, C. Le Bas, M. Jamagne, D. King & L. Montanarella (eds). Soil Geographical Database for Eurasia & The Mediterranean: Instructions Guide for Elaboration at scale 1:1,000,000, Version 4.0. European Soil Bureau Research Report No.8, EUR 20422 EN, (2003), 64pp. Office for Official Publications of the European Communities, Luxembourg.

Lillesand, T M and R W Kiefer, 1994. Remote sensing and image interpretation. John Wiley & Sons, New York, 3rd edition.

Ma, Z. and R L Redmond, 1995. Tau coefficients for accuracy assessment of classification of remote sensing data. Photogrammetric Engineering & Remote Sensing, 61(4):435-439.

Minasny, B. and A.B. McBratney, 2000. FuzME version 2.1. Australian Centre for precision Agriculture, University of Sydney, NSW 2006. http://www.usyd.edu.au/su/agric/acpa/fkme/FkME.html (accessed 20 June 2004).

Moellering, H., 1987. A draft proposed standard for digital cartographic data. Report no. 8. USA National Committee for Digital Cartographic Standards, American Congress on Surveying and Mapping.

Næsset, E., 1996. Use of the weighted Kappa coefficient in classification error assessment of thematic maps. International Journal of Geographical Information Systems, 10(5):591–603.

Odeh, I.O.A., A.B. McBratney and D.J.Chittleborough, 1992. Soil pattern recognition with fuzzy c-means: Application to classification and soil-landform interrelationship. Soil Sci Soc Am J 56: 505-516.

Pebesma, E.J. and G.C. Wesseling. Gstat, a program for geostatistical modelling, prediction and simulation. Computer Geoscri. 24: 17-31.
Radoševic, N., 1979. Pre-war military map 1:100 000 (1:50 000) and todays triangulation. VGL, 129-148.
Raiffa, H. and R. Schlaifer, 2000. Applied Statistical Decision Theory. Wiley, New York, ISBN: 0-471-38349-X. 356 p.
RIVM, 1997. Nationale milieuverkenningen 1997-2020. RIVM, Bilthoven, The Netherlands.
Rubin, D. B. (1987). Multiple Imputation for Non-response in Surveys. New York: Wiley.
De Bruin, S. and A. Stein, 1998. Soil-landscape modelling using fuzzy c-means clustering of attribute data from a Digital Elevenation Model (DEM). Geoderma 83: 17-33.
Van Engelen, V.W.P. and T.T. Wen (eds.), 1995. Global and national soils and terrain digital databases (SOTER). Procedures Manual (revised edition). UNEP-ISSS-ISRIC-FAO. Wageningen, The Netherlands.
van Reeuwijk, L.P. and V.J.G. Houba, 1998. Guidelines for Quality Management in Soil and Plant Laboratories. (FAO Soils Bulletin - 74). Food and Agriculture Organization of the United Nations, Rome.
Walter, V. & Fritsch, D,. 1999. Matching Spatial Data Sets: a Statistical Approach, International Journal of Geographical Information Science 13(5), 445–473.
Van Ranst E., Thomasson A.J., Daroussin J., Hollis J.M., Jones R.J.A., Jamagne M., King D. & Vanmechelen L. (1995). Elaboration of an extended knowledge database to interpret the 1:1,000,000 EU Soil Map for environmental purposes. In: European Land Information Systems for Agro-environmental Monitoring. D. King, R.J.A. Jones and A.J. Thomasson (eds.). EUR 16232 EN, 71-84. Office for Official Publications of the European Communities, Luxembourg.
Webster, R. and M.A. Oliver, 1990. Statistical methods in soil and land resource survey. Oxford University Press, Oxford.
Wösten, J.H.M., P.A. Finke and M.J.W. Jansen, 1995. Comparison of class- and continuous pedotransfer functions to generate soil hydraulic characteristics. Geoderma 66: 227-237.
Zhang, J., Goodchild, M. 2002. Uncertainty in Geographical Information. Taylor & Francis. ISBN: 0415243343.

Table 1 Similarity relations between components of uncertainty and data quality (Fisher, 2003).

Uncertainty	Data Quality
Error	- Positional Accuracy - Attribute Accuracy - Completeness
Vagueness, Discord, Ambiguity	Semantic Accuracy
Error, Discord, Vagueness, Ambiguity	Currency
Discord	Logical consistency
?	Lineage

Table 2 Impacts of uncertain data on different groups of users. U=uncertainty.

User	Positive aspect	Negative aspect	User profits by:
Researcher	Window of improvement	Damaged public image	- quantification of U - identification of sources of U - minimising U
Policy maker	Window of decision	Wrong decisions Damaged public image	- quantification of U - minimising U - deciding in the presence of U - U or risk communication
Stakeholder	Window of argumentation	Wrong decisions	- quantification of U - deciding in the presence of U - U or risk communication

Table 3 Confusion matrix with proportions of observations within *c* mapped classes *i* and ground truth classes *j*.

		Ground truth class *j*			
		j=1	…	*j*=*c*	Total
Mapped class *i*	*i*=1	$p_{1,1}$	…	$p_{1,c}$	$p_{1+} = \sum_{j=1}^{c} p_{1j}$
	…	…	$p_{i,j}$	…	$p_{i+} = \sum_{j=1}^{c} p_{ij}$
	i=*c*	$p_{c,1}$	…	$p_{c,c}$	$p_{c+} = \sum_{j=1}^{c} p_{cj}$
	Total	$p_{+1} = \sum_{i=1}^{c} p_{i1}$	$p_{+j} = \sum_{i=1}^{c} p_{ij}$	$p_{+c} = \sum_{i=1}^{c} p_{ic}$	$\sum_{i=1}^{c} \sum_{j=1}^{c} p_{ij} = 1$

Figure captions

Figure 1 Evolution of map quality parameters (ground water table class map 1:50,000, sheet 27 East, The Netherlands) for 10 consecutive years after an update, estimated with *n=52-73* monitoring wells.

Figure 2 Logical inconsistency in map generalisation

Figure 3 WRB major soil group (source: http://eusoils.jrc.it/msapps/Soil/SoilDB/SoilDB.phtml)

Figure 4 Uncertainty in the data collection and the data application domains. Dotted lines indicate interactions between the domains.

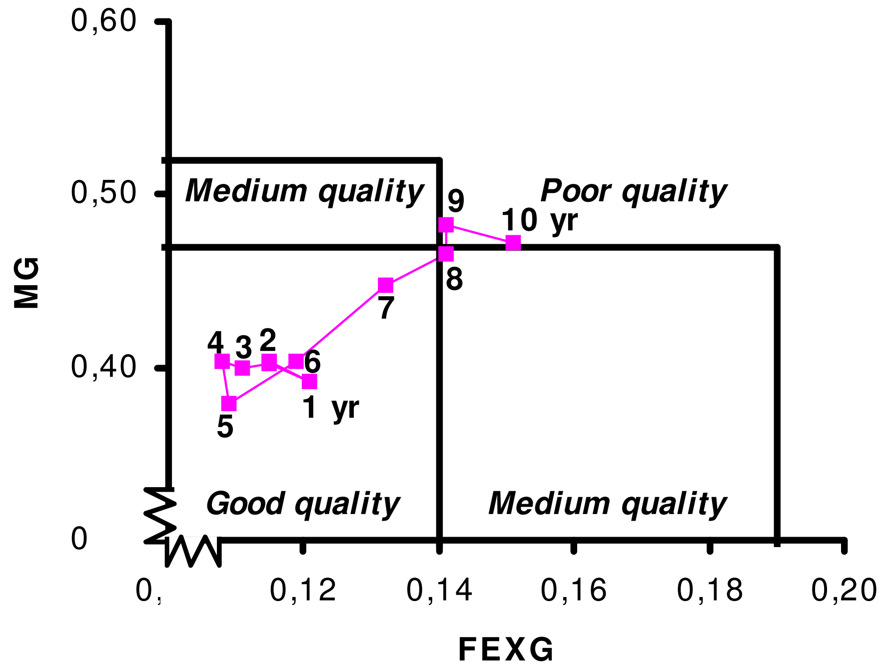

Figure 1 Evolution of map quality parameters (ground water table class map 1:50,000, sheet 27 East, The Netherlands) for 10 consecutive years after an update, estimated with *n*=52-73 monitoring wells.

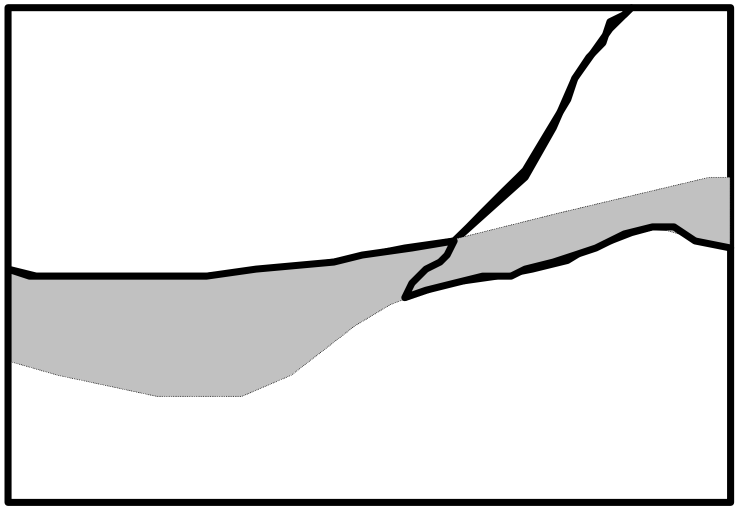

Figure 2 Logical inconsistency in map generalisation.

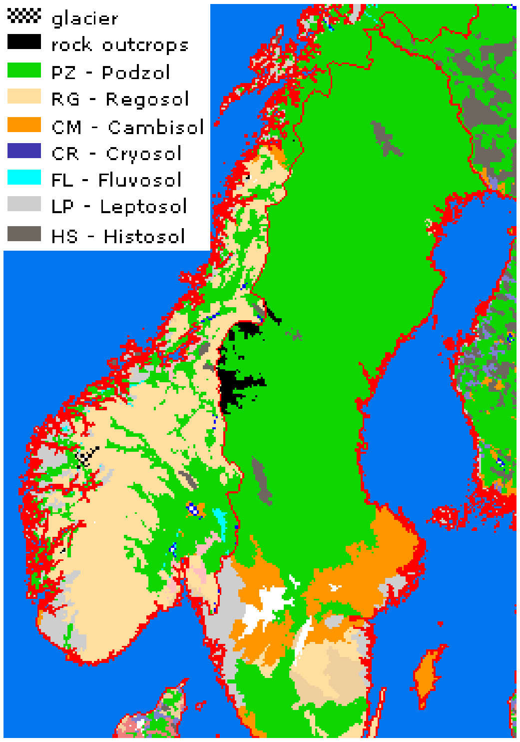

Figure 3 WRB major soil group (source: http://eusoils.jrc.it/msapps/Soil/SoilDB/SoilDB.phtml).

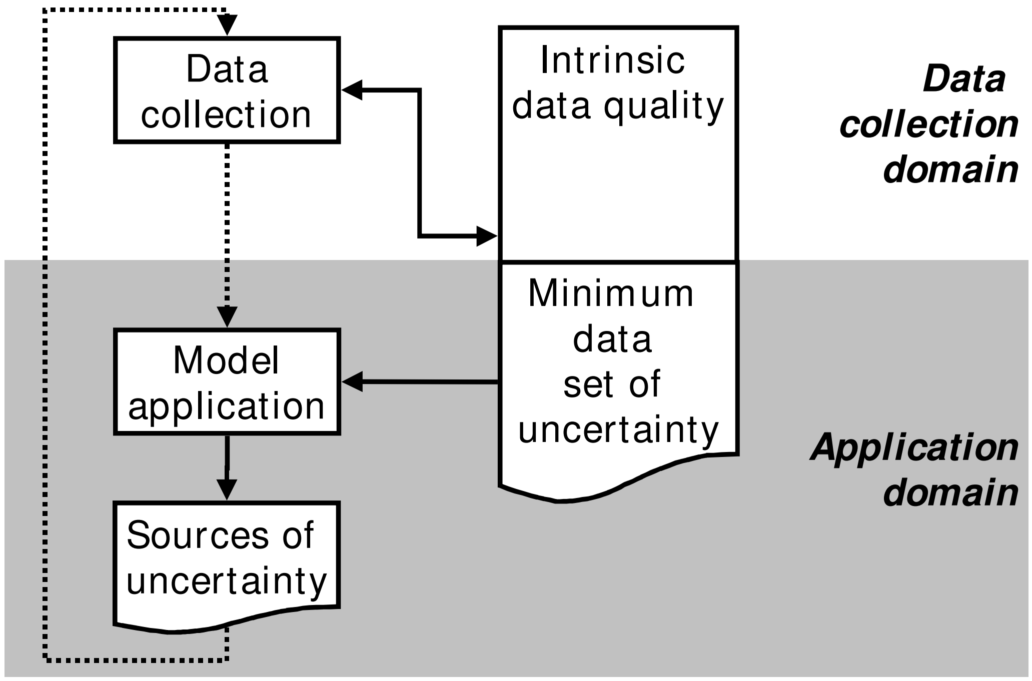

Figure 4 Uncertainty in the data collection and the data application domains. Dotted lines indicate interactions between the domains.